CLAIMS

What is claimed is:

1. A method for evaluating and outputting a final clustering solution for a plurality of multi-dimensional data records, said data records having multiple, heterogeneous feature spaces represented by feature vectors, said method comprising:

defining a distortion between two feature vectors as a weighted sum of distortion measures on components of said feature vector;

clustering said multi-dimensional data records into k-clusters using a “convex programming” formulation; and

selecting feature weights of said feature vectors.

2. The method according to claim 1, wherein said selecting of feature weights are optimized by an "objective" function to produce said solution of a final clustering that simultaneously minimizes average intra-cluster dispersion and maximizes average inter-cluster dispersion along all said feature spaces.

3. The method according to claim 1, wherein said clustering includes initially applying a local minima of said clustering.

4. The method of claim 1, wherein said clustering comprises a k-means clustering algorithm.

5. The method of claim 2, wherein said minimizing distortion of individual clusters includes taking said data records and iteratively determining *Voronoi* partitions until said “objective” function, between two successive iterations, is less than a specified threshold.

6. The method of claim 1, wherein said clustering comprises analyzing word data, and said feature vectors comprise multiple-word frequencies of said data records.

7. The method of claim 1, wherein said clustering comprises analyzing data records having numerical and categorical attributes, and said feature vectors comprise linearly-scaled numerical attributes and each q-ary categorical feature using a 1-in-q representation of said data records.

8. A method for evaluating and outputting a clustering solution for a plurality of multi-dimensional data records, said data records having multiple, heterogeneous feature spaces represented by feature vectors, said method comprising:

defining a distortion between two said feature vectors as a weighted sum of distortion measures on components of said feature vector;

clustering said multi-dimensional data records into k-clusters using a “convex programming” formulation of a generalized k-means clustering function; and

selecting optimal feature weights of said feature vectors by an "objective" function to produce said solution of a final clustering that simultaneously minimizes average intra-cluster dispersion and maximizes average inter-cluster dispersion along all said feature spaces.

9. The method of claim 8, wherein said clustering includes initially applying a local minima of said clustering.

10. The method of claim 8, wherein said minimizing distortion of individual clusters includes taking said data records and iteratively determining *Voronoi* partitions until said “objective” function, between two successive iterations, is less than a specified threshold.

11. The method of claim 8, wherein said clustering comprises analyzing word data, and said feature vectors comprise multiple-word frequencies of said data records.

12. The method of claim 8, wherein said clustering comprises analyzing data records having numerical and categorical attributes, and said feature vectors comprise linearly-scaled numerical attributes and each q-ary categorical feature using a 1-in-q representation of said data records.

13. A computer system for data mining and outputting a final clustering solution, wherein said system includes a memory for storing a database having a plurality of multi-dimensional data records, each having multiple, heterogeneous feature spaces represented by feature vectors, said system including a processor for executing instructions comprising:

defining a distortion between two feature vectors as a weighted sum of distortion measures on components of said feature vector;

clustering said multi-dimensional data records into k-clusters using a “convex programming” formulation; and

selecting feature weights of said feature vectors.

14. The system of claim 13, wherein said instruction for selecting of said feature weights are optimized by implementing an "objective" function to produce said solution of a final clustering that simultaneously minimizes average intra-cluster dispersion and maximizes average inter-cluster dispersion along all said feature spaces.

15. The system of claim 13, wherein said instruction of said clustering includes an instruction for initially applying a local minima of said clustering.

16. The system of claim 13, wherein said instruction for clustering includes instructions for implementing a k-means clustering algorithm.

17. The system of claim 14, wherein said instruction for minimizing distortion of individual clusters includes taking said data records and iteratively determining *Voronoi* partitions until said "objective" function, between two successive iterations, is less than a specified threshold.

18. The system of claim 13, wherein said instruction for clustering includes instructions for analyzing word data.

19. The system of claim 13, wherein said instruction for clustering includes instructions for analyzing data records having numerical and categorical attributes.

20. A program storage device readable by machine, tangibly embodying a program of instructions executable by said machine to perform a method for evaluating and outputting a final clustering solution from a set of data records having multiple, heterogeneous feature spaces represented as feature vectors, said method comprising:

defining a distortion between two feature vectors as a weighted sum of distortion measures on components of said feature vector;

clustering said multi-dimensional data records into k-clusters using a “convex programming” formulation; and

selecting feature weights of said feature vectors.

21. The device of claim 20, wherein said selecting of feature weights are optimized by an "objective" function to produce said solution of a final clustering that simultaneously minimizes average intra-cluster dispersion and maximizes average inter-cluster dispersion along all said feature spaces.

22. The device of claim 20, wherein said clustering includes initially applying a local minima of said clustering.

23. The device of claim 20, wherein said clustering comprises a k-means clustering algorithm.

24. The device of claim 21, wherein said minimizing distortion of individual clusters includes taking said data records and iteratively determining *Voronoi* partitions until said “objective” function, between two successive iterations, is less than a specified threshold.

25. The device of claim 20, wherein said clustering comprises analyzing word data, and said feature vectors comprise multiple-word frequencies of said data records.

26. The device of claim 20, wherein said clustering comprises analyzing data records having numerical and categorical attributes, and said feature vectors comprise linearly-scaled numerical attributes and each q-ary categorical feature using a 1-in-q representation of said data records.